AISSMS

INSTITUTE OF INFORMATION TECHNOLOGY

ADDING VALUE TO ENGINEERING

A PRELIMENERY REPORT ON

"CRIME PREDICTION USING SENTIMENTAL ANALYSIS "

SUBMITTED TO THE SAVITRIBAI PHULE UNIVERSITY, PUNE
IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

OF

BACHELOR OF ENGINEERING (COMPUTER ENGINEERING)

SUBMITTED BY

Mr. BEZAN NANAVATTI	Exam No. 72024009D
Ms. KAVITA THAKUR	Exam No. 72023919C
Mr. SUNIT LOHADE	Exam No. 72023953C
Mr. RUTVIK KORADE	Exam No. 72037924F

Under The Guidance of

Dr. Minal Zope

DEPARTMENT OF COMPUTER ENGINEERING

ALL INDIA SHRI SHIVAJI MEOMRIAL SOCIETY'S

INSTITUTE OF INFORMATION TECHNOLOGY

KENNEDY ROAD, NEAR R.T.O. PUNE-411001

SAVITRIBAI PULE PUNE UNIVERSITY

2020-2021

AISSMS
INSTITUTE OF INFORMATION TECHNOLOGY
(IOIT)
ADDING VALUE TO ENGINEERING
An Autonomous Institute Affiliated to Savitribai Phule Pune University
Approved by AICTE, New Delhi and Recognised by Govt. of Maharashtra
Accredited by NAAC with "A+" Grade | NBA - 5 UG Programmes

CERTIFICATE

This is to certify that the project report entitles

"CRIME PREDICTION USING SENTIMENT ANALYSIS "

Submitted by

Mr. BEZAN NANAVATTI	Exam No. 72024009D
Ms. KAVITA THAKUR	Exam No. 72023919C
Mr. SUNIT LOHADE	Exam No. 72023953C
Mr. RUTVIK KORADE	Exam No. 72037924F

is a bonafide student of this institute and the work has been carried out by him/her under the supervision of **Dr. M. Zope** and it is approved for the partial fulfillment of the requirement of Savitribai Phule Pune University, for the award of the degrees of **Bachelor of Engineering** (Computer Engineering).

(Dr. M. Zope)
Guide
Department of Computer Engineering

(Dr. S.N. Zaware)
Head
Department of Computer Engineering

(Dr. P.B. Mane)
Principal
AISSMS Institute of Information Technology, Pune-411001

Place: PUNE
Date:

ACKNOWLEDGMENT

It gives us immense pleasure in presenting the project report on **"CRIME PREDICTION USING SENTIMENT ANALYSIS"** The success and the outcome of this report required a lot of guidance. We are very grateful to our guide **Dr. M. Zope** who has provided expertise and encouragement. We thank sir who provided vision and knowledge that was very helpful throughout the research. All that we have done is only due to the great guidance. We also express our gratitude to **Dr. S.N. Zaware** Head of Computer Engineering Department, AISSMS's Institute of Information Technology, for the valuable support.

We would also like to extend our sincere thanks to Principal Dr.P.B. Mane, for his dynamic and valuable guidance throughout the project and providing the necessary facilities that helped us to complete our dissertation work.

Mr. Bezan Nanavatti
Ms. Kavita Thakur
Mr. Rutvik Korade
Mr. Sunit Lohade

ABSTRACT

Crime is one of the biggest and dominating problem in our society and its prevention is an important. task. Daily there are huge numbers of crimes committed frequently. This require keeping track of all the crimes and maintaining a database for same which may be used for future reference. The current problem faced are maintaining of proper dataset of crime and analyzing this data to help in predicting and solving crimes in future. The objective of this project is to analyze dataset which consist of numerous crimes and predicting the type of crime which may happen in future depending upon various conditions. Before training of the model data preprocessing will be done following this feature selection and scaling will be done so that accuracy obtain will be high.

The K- Nearest Neighbor (KNN) classification and various other algorithms will be tested for crime prediction and one with better accuracy will be used for training. Visualization of dataset will be done in terms of graphical representation of many cases for example total number of crime reports of an area per year along with the sentiment or emotions that area exhibits. The sole purpose of this project is to give a jest idea of how machine learning can be used by the law enforcement agencies to detect, predict and solve crimes at a much faster rate and thus reduces the crime rate. It not restricted to India, this can be used in other states or countries depending upon the availability of the dataset.

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Overview

Crime is one of the biggest and dominating problem in our society and its prevention is an important. task. A huge number of crimes are committed frequently. This requires keeping track of all the crimes and maintaining a database for same which may be used for future reference. The current problem faced are maintaining of proper dataset of crime and analyzing this data to help in predicting and solving crimes in future.

1.2 Motivation

The objective of this project is to analyze dataset which consist of numerous crimes and predicting the type of crime which may happen in future depending upon various conditions. Before training of the model data preprocessing will be done following this feature selection and scaling will be done so that accuracy obtain will be high. Sentiment analysis (SA) is widely used today in many areas such as crime detection (security intelligence) to detect potential security threats in real-time using social media platforms such as Twitter. In this day and age, a lot of people tend to gauge situations more frequently based on emotions rather than solid facts To be able to exhibit the same, we use a technique called sentimental analysis to understand the public's viewpoints and feelings regarding a certain situation or occurrence Due to the exponential increase in crime rates occurring now, we need a system not only based on historical events, but to predict the heinousness of crimes that may occur in the future We believe that using this model of crime prediction using sentiment analysis, we may be able to overcome at least a noticeable amount of upcoming difficulties regarding crime.

1.3 Problem Definition and Objectives

According to criminal investigators, the behaviour displayed on social media can help identify potential criminal intent, which is useful in crime prevention. Social media platforms, however, serve billions of users, making manual social media policing almost impossible, but with the use of modern data analytics tools law enforcement can scour social media platforms for precursors to criminal intent. Modern machine learning tools embedded into data analytics platforms can analyze billions of comments and other social media posts to identify criminal intent, allowing authorities to apprehend these individuals and prevent crime. Thus, our

objective is to create such a model that can provide us with the information and statistics we need to predict the crime and possibly help control and avoid it.

1.4 Project Planning and Scheduling

• Project planning is part of project management, which relates to the use of schedules such as Gantt charts to plan and subsequently report progress within the project environment.
• Initially, the project scope is defined and the appropriate methods for completing the project are determined. Following this step, the durations for the various tasks necessary to complete the work are listed and grouped into a work breakdown structure.
• Project planning is often used to organize different areas of a project, including project plans, workloads and the management of teams and individuals.

1.5 Project Plan

Planning before any activity is very important. And if it is planned well enough, then success is almost guaranteed. A standard Project Management System has six major modules of Admin, Manage Application, Test Management, Process Management, Manage Comment, Reports. We analyzed the overall complexity of each of these modules and it was found that the project will require approximately 14 to 28 weeks to approach completion, so we planned accordingly. We decided to follow the SDLC i.e. Software Development Life Cycle while planning various phases of our project. This method consists of following activities:
1. Determination of system requirements
2. System Analysis
3. Design of system
4. Development of software
5. System Testing
6. Implementation and Evaluation

1. Project Scheduling:
   • Project Scheduling is the culmination of a planning activity that is primary component of software project management.
   • When combined with estimation methods and risk analysis, scheduling,

establishes a road map for the project management.

• Scheduling begins with the process composition. The characteristics of the project are used to adapt an appropriate task set for the work to be done.

• The task network is used to compute the critical project path, a time line chart and a variety of project information.

• When creating a software project schedule, the planner begins with a set of tasks. If automated tools are used, the work breakdown is input as a task network or task outline. Effort, duration, and start date are then input for each task. In addition, tasks may be assigned to specific individuals.

CHAPTER 2

LITERATURE SURVEY

1. Pratibha, A. Gahalot, Uprant, S. Dhiman and L. Chouhan, "Crime Prediction and Analysis," 2nd International Conference on Data, Engineering and Applications (IDEA), 2020, pp. 1-6, doi: 10.1109/IDEA49133.2020.9170731. In Recent time, many researchers have conducted experiments to predict the crimes using various machine learning methods.
2. Kumar, A. Verma, G. Shinde, Y. Sukhdeve and N.Lal, "Crime Prediction Using K-Nearest Neighboring Algorithm," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-4, doi: 10.1109/ ic-ETITE47903.2020.155. We observe crime rates by the KNN prediction method. It will predict, tentatively, the type of crime, when, where and at what time it may take place.
3. N. H. M. Shamsuddin, N. A. Ali and R. Alwee, "An overview on crime prediction methods," 2017 6th ICT International Student Project Conference (ICT-ISPC), 2017, pp. 1-5, doi: 10.1109/ICTISPC.2017.8075335. The performance of prediction models can be evaluated using a variety of different prediction methods such as support vector machine, multivariate time series and artificial neural network.
4. W. Safat, S. Asghar and S. A. Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques," in IEEE Access, vol. 9, pp. 70080-70094, 2021, doi:10.1109/ACCESS.2021.3078117. This study applied different machine learning algorithms, namely, the logistic regression, support vector machine (SVM), Naïve Bayes, k-nearest neighbors (KNN), decision tree, multilayer perceptron (MLP), random forest, and eXtreme Gradient Boosting (XGBoost), and time series analysis by long-short term memory (LSTM) and autoregressive integrated moving average (ARIMA) model to better fit the crime data.
5. S. G. Krishnendu, P. P. Lakshmi and L. Nitha, "Crime Analysis and Prediction using Optimized KMeans Algorithm," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 915-918
6. J. Kiran and K. Kaishveen., "Prediction Analysis of Crime in India Using a Hybrid Clustering Approach," 2018 2nd International Conference on ISMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference on, 2018, pp.520-523, doi: 10.1109/ISMAC.2018.8653744.
7. X. Chen, Y. Cho and S. Y. Jang, "Crime prediction using Twitter sentiment and weather,"2015 Systems and Information Engineering Design Symposium, 2015,

pp. 63-68, doi: 10.1109/ SIEDS.2015.7117012. The aim is to predict the time and location in which a specific type of crime will occur.
8. M. Al Boni and M. S. Gerber, "Predicting crime with routine activity patterns inferred from social media," 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016, pp. 001233-001238, doi: 10.1109/SMC.2016.7844410 Prior work in statistical crime prediction has not investigated micro-level movement patterns of individuals in the area of interest.
9. Gerber, Matthew. (2014). Predicting Crime Using Twitter and Kernel Density Estimation. Decision Support Systems. 61. 10.1016/j.dss.2014.02.003. This article presents research investigating the use of spatiotemporally tagged tweets for crime prediction.
10. X. Chen, Y. Cho and S. Y. Jang, "Crime prediction using Twitter sentiment and weather," 2015 Systems and Information Engineering Design Symposium, 2015, pp. 63-68, doi: 10.1109/SIEDS.2015.7117012.The aim is to predict the time and location in which a specific type of crime will occur.

CHAPTER 3

SOFTWARE REQUIREMENTS SPECIFICATION

3.1 Assumptions and Dependencies

Assumptions

1. 1. The system is designed for English language and the input text for summarization must be in this language.

   2. The user must have Web Browser.

   3. The web app requires a proper network connection.

Dependencies

NLTK, Numpy, Pandas, Tkinter, Geopandas, Seaborn, Matplotlib, Shapefile, Shapely, Pyinstaller, Textblob, Pyogrio, Sklearn, Utils, os, Pyshp, Mapclassify, etc.

3.2 Non-Functional Requirements

3.2.1 Usability

- The system must be easy to use. The user must get the summarized text with one button press if possible.
  - Being timesaving is one of he software feature.
  - It must also should be user friendly for everyone using it.

3.2.2 Reliability

- This software is being developed with NLP. So, there is not much data to measure.

  - The user provided data is compared with the result for checking accuracy.
  - With recent NLP techniques, user gained data should be used efficiently for generating summary. The maintenance period should be as short as possible.

  - The users must be able to use the program easily, so maintenance should not be a problem.

3.2.3 Performance

- Calculation and response time must be small, as it must be timesaving.
  • The process of summarizing should not take more than 30 seconds in order to summarize 1 page long document. The server must be as fast as possible.

  • Output must be generated fast, as multiple user might access the sight con- currently.

  • 1-minute delay of response time should be allowed. The certain limit is acceptable. It can be given to user with an appropriate message.

3.2.4 Supportability

- The system is based Python and NLP. So, problem on server side requires prior knowledge of the filed.

  • Client side problems requires web technology and networking knowledge.

3.2.5 Software Requirements

- Programming Language: Python
  • IDE: Jupyter Notebook

3.2.6 Hardware Requirements

- Processor: Intel i3 and above
  • Operating System: Windows 10
  • Ram: 4GB and above
  .

3.2.7 Web Development Kit

- IDE: Angular and Flask

3.3 Analysis Models: SDLC Model to be applied

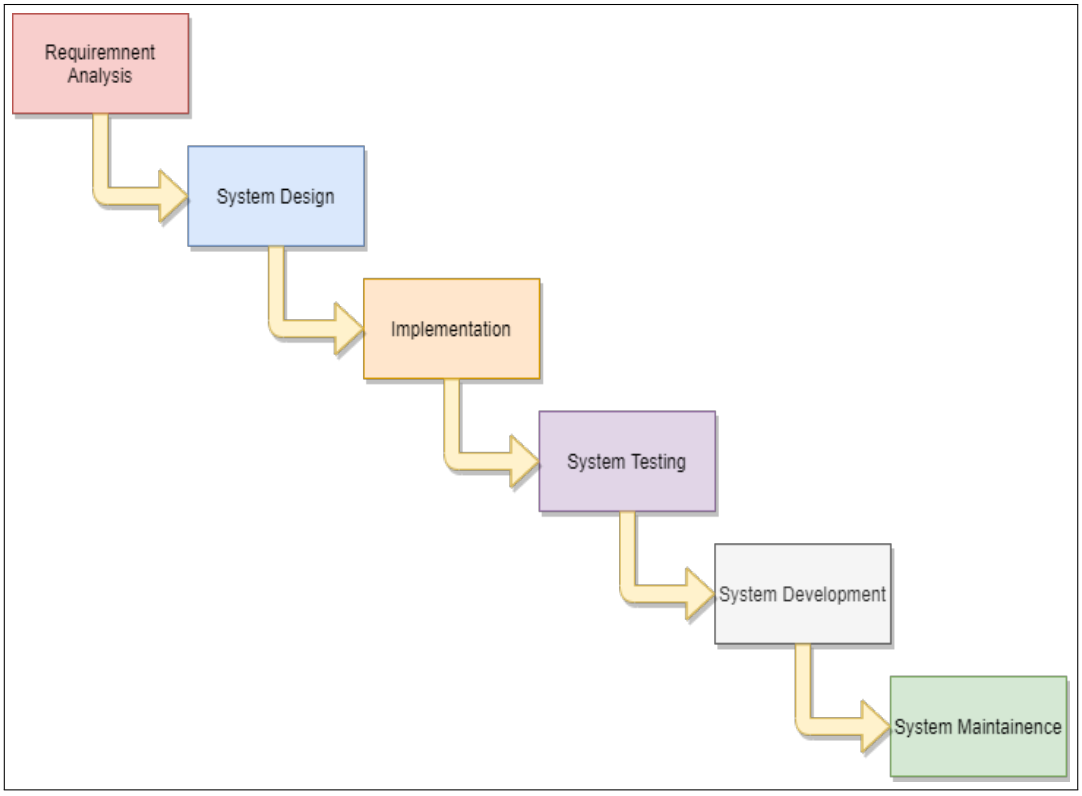


Figure 3.1: Waterfall Model

SDLC Waterfall model is being depicted by our system.

- The initial stage is requirement analysis stage here the data is being gathered which is to be provided as an input to the system.
- Second stage is the design stage where all the data is being formatted into a particular matrix. The scores are used to generate the matrix.
- Third stage is the coding stage in which the system performs its main functionality of mapping of the classes and getting the exact prototype.
- Testing is done in the fourth phase inorder to test the word with the probabilistic label.
- In the maintenance phase it's the last phase wherein the system has to depict and maintain the probabilistic labels of the respective words.

3.4 System Implementation Plan

In the system plan implementation the input can be in text, .csv file format. This data entered is first cleaned and then processed using NLP and text cleaning, sorting algorithms. After this step based on the scores generated sentences are ranked for output summary generation.

CHAPTER 4

SYSTEM DESIGN

4.1 System Architecture

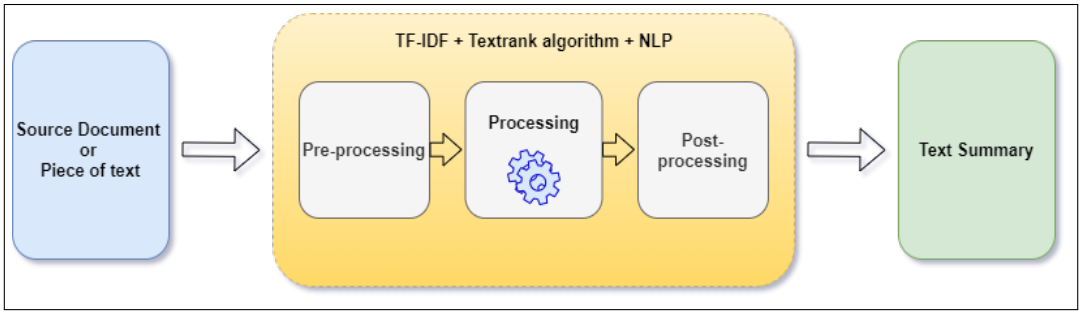


Figure 4.1: System Architecture

4.2 Data Flow (DFD) Diagrams

4.2.1 DFD Level 0 Diagram

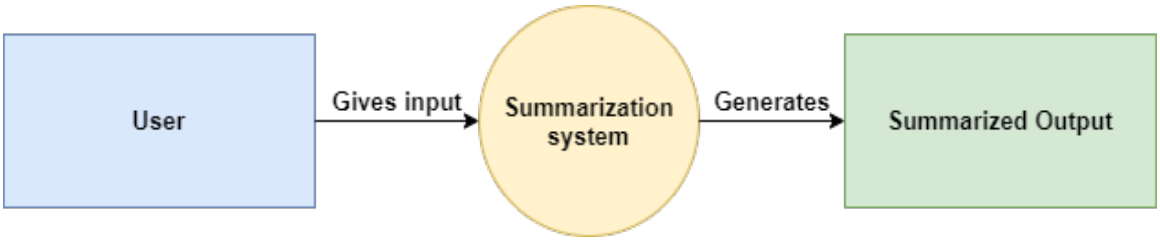


Figure 4.2: DFD Level 0 diagram of Text Summarization System

Level 0 DFD proposed system show the actors input and output data from the system.

As shown in Figure 4.2, there are three parts, i.e. user gives input to the summarization system in the form of text document or a text paragraph for getting a summarized output.

4.2.2 DFD Level 1 Diagram

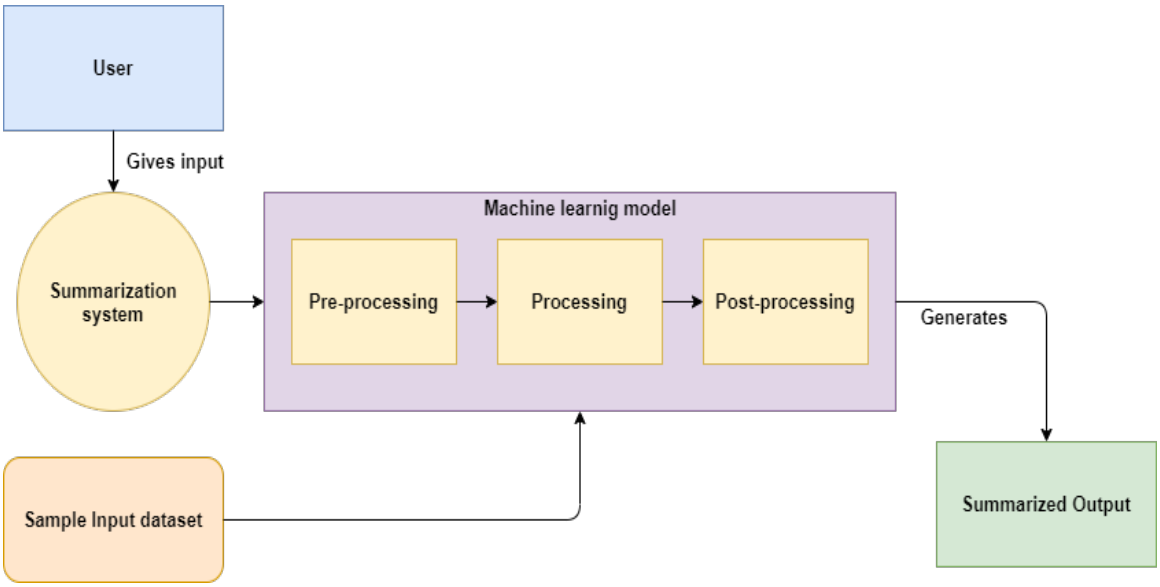


Figure 4.3: DFD Level 1 diagram of Text Summarization System

Level 1 DFD aims to give us a deeper look of the system. Larger processes are broken down into sub-processes.
As shown in Figure 4.3, the user has to give input to proceed further. The system will clean and extract the feature, i.e. removing stops words, matrix generation, ranking of sentences etc. and will give the output. Here textrank algorithm is used with some NLP libraries and is also trained on the BBC news dataset.

4.3 UML Diagrams

4.3.1 Activity diagram

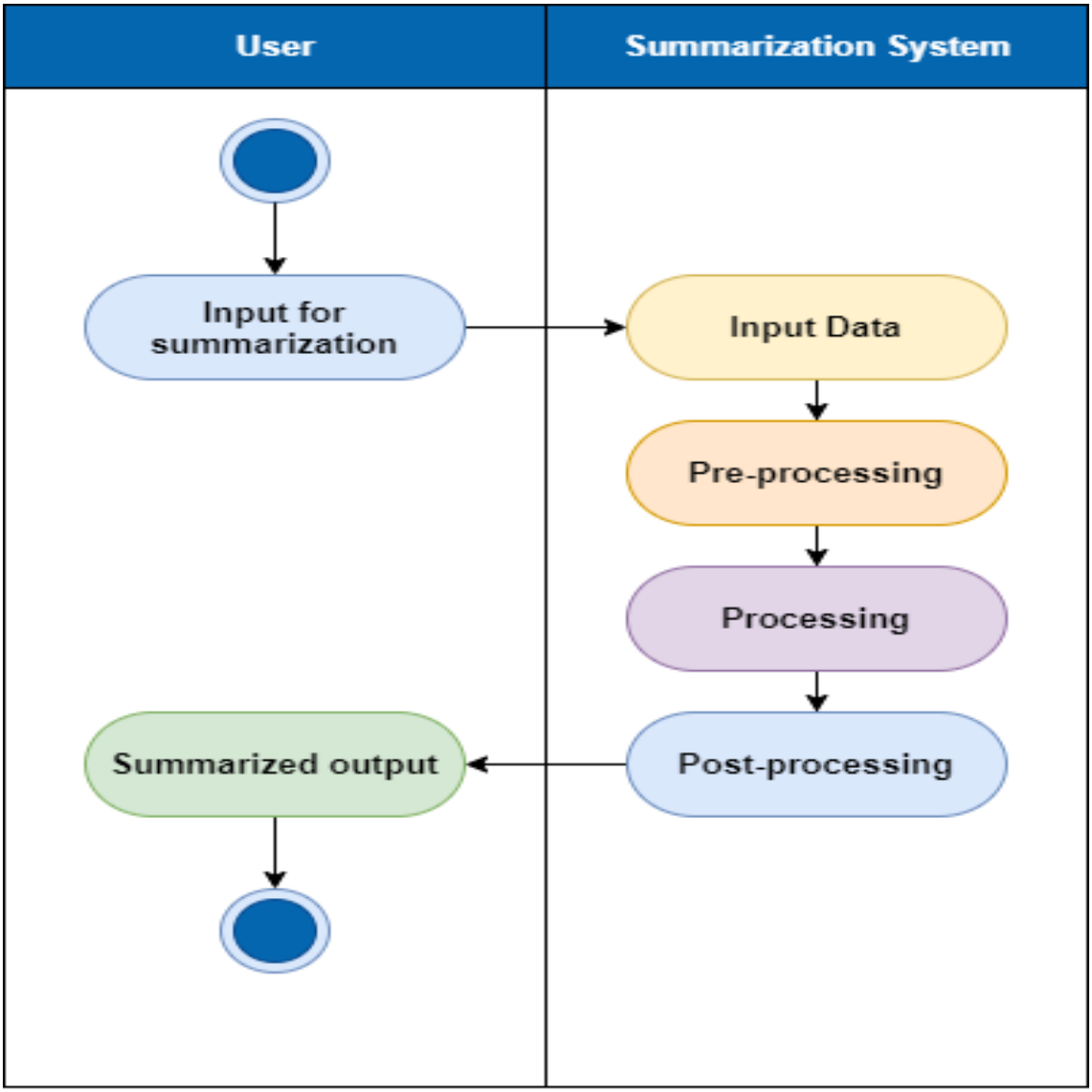


Figure 4.4: Activity diagram of Text Summarization System

- Activity Diagram is a behavioral diagram presenting the actors their functions performed.
- They also include the swim lanes and the forks and joins.
- The represent the individual lane as their entire activities and the functionality carried out by that particular actor in the respective lanes.

- Input is being provided by the user and the output is being given back to the user.

4.3.2 Sequence Diagram

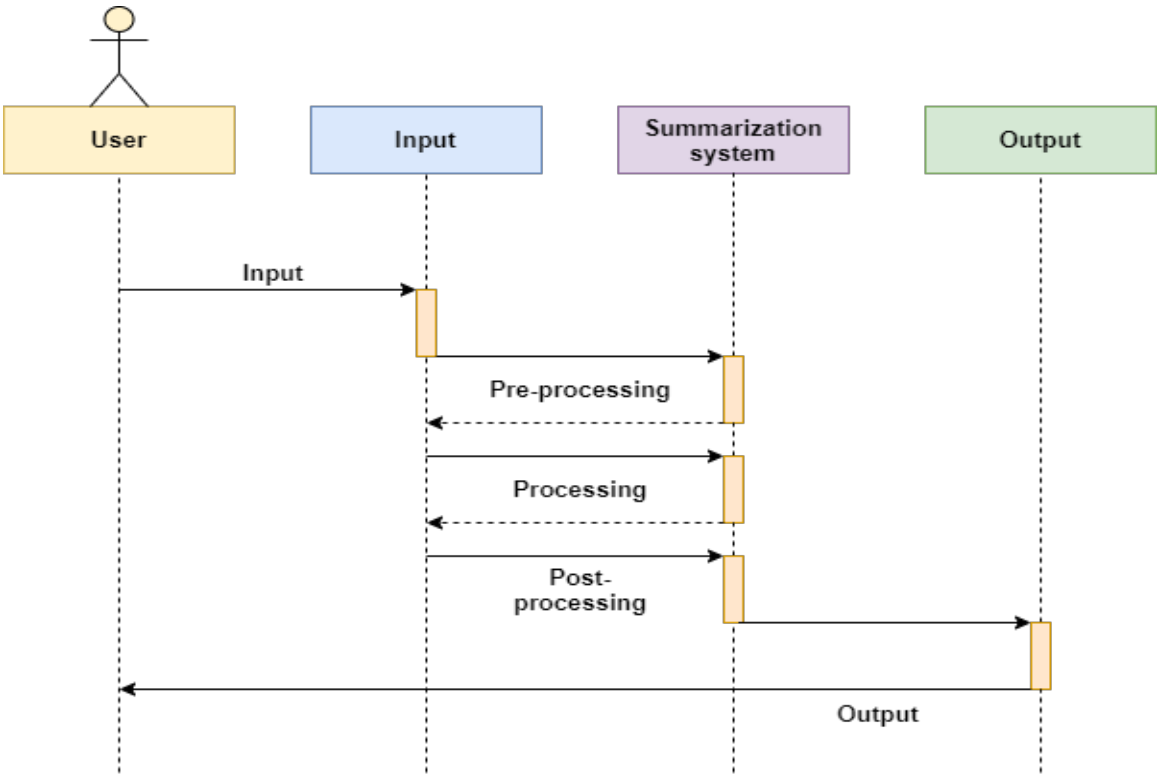


Figure 4.5: Sequence diagram of Text Summarization System

In the above Figure 4.5 there are different actors:

- Actors
  - User
  - Input
  - Summarization System
  - Output
- Different roles played by the different actors are been specified separately in their individual lanes.
- Each actor has number of timelines for the activity (Function) to be pre-formed.

4.3.3 Usecase Diagram

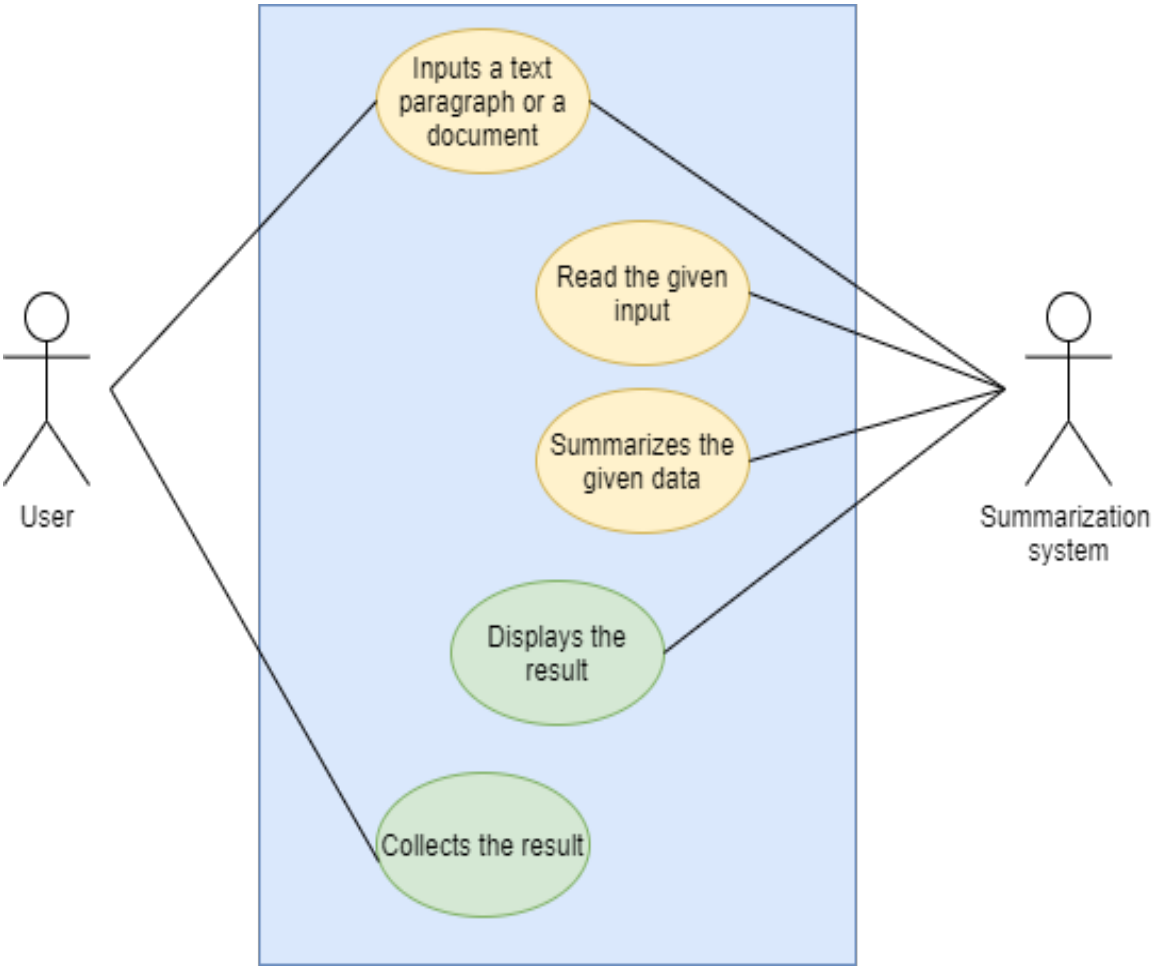


Figure 4.6: Usecase diagram of Text Summarization System

Use Case diagram is used for representing the problem statement that is the actors in it, their functionality in an behavioral manner. They are useful when the system is to be in the programmatic execution.
In the above Figure 4.6 there are different actors:

- Actors
  - User
  - Summarization System
- Functional Blocks include the stepwise execution of the entire system that is the form of different procedures (functionality).
- Input to the every function in this diagram is the output from the previous state.

4.3.4 System Architecture

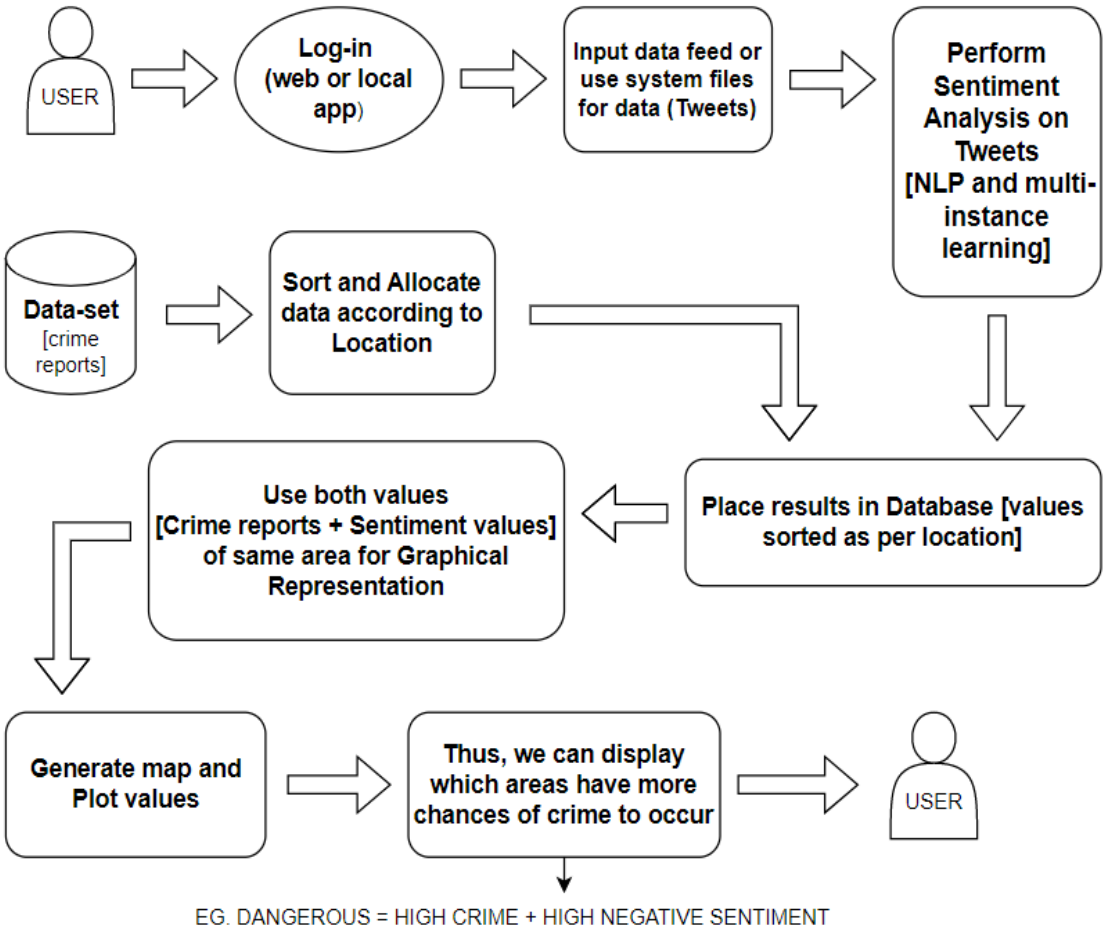


Figure 4.7: System architecture and working

CHAPTER 5

PROJECT PLAN

5.1 Team Organization

Our team consists of four members and the members of our team are allocated certain tasks to complete

5.1.1 Analysis

Members :

Bezan Nanavatti, Kavita Thakur, Rutvik Korade, Sunit Lohade

5.1.2 Design Coding

Members :

Bezan Nanavatti

5.1.3 Documentation

Members :

Bezan Nanavatti, Kavita Thakur, Rutvik Korade, Sunit Lohade

5.2 Risk Management

Project Risk Management involves conducting risk management planning, engaging in risk identification, completing risk analysis, creating a risk response action plan, and monitoring and controlling risk on a project. Project Risk Management is a continuous process to be engaged in throughout the entire project. A key point to remember is that risk is not always bad. There are opportunities and there are threats. The opportunities are the good risks. The treats are the bad risks.

5.2.1 Risk Identification

Risk Identification is a systematic attempt to specify threats to the project plan. By identifying the known and predictable risks, the project manager takes a first step towards avoiding them when possible and controlling them when necessary One method of identifying risks is to create a risk item checklist. The checklist can be used for risk identification and focuses on some subset of known and predictable risks in the following subcategories:

• Product Size –risks associated with the overall size of the software to be built or modified.
• Business impact –risks associated with constraints imposed by management.
• Customer Characteristics – risks associated with the sophistication of the customer and the developer's ability to communicate with the customer in a timely manner.
• Process Definition – risks associated with the degree to which the software process has been defined and is followed by the development organization.
• Development Environment – risks about the availability and quality of the tools to be used to build the project.
• Technology to be built – risk on the complexity of the system to be built and the newness of the technology.

• Staff size and experience – risks with the overall technical and project experience of the software engineers who will handle the work.

5.2.2 Risk Analysis

5.2.3 Overview of Risk Mitigation, Monitoring, Management

Table 5.1: Risk 1 Mitigation, Monitoring, Management

Risk ID	1
Risk Description	Internet Availability
Category	Requirements
Probability	Less
Impact	High
Response	Mitigation
Strategy	Wi-Fi, LAN, Router
Risk Status	Identified

Table 5.2: Risk 2 Mitigation, Monitoring, Management

Risk ID	2
Risk Description	Need of Good Processing Power
Category	Requirements
Probability	Medium
Impact	High
Response	Mitigation
Strategy	Dataset Pre-processing
Risk Status	Identified

5.3 Project Schedule

5.3.1 Project Task Set

Major aspects of the project:

- Task 1: Correctness
- Task 2: Availability
- Task 3: Integrity

5.3.2 Timeline Chart

Table 5.3: System Implementation Plan

SR NO	DURATION	ACTIVITY PERFORMED
1	July second week	Topic Finalization
2	3rd and 4th week of July	Understanding of Base Paper
3	1st and 2nd week of August	Literature Survey
4	3rd and 4th week of August	System Architecture Design Completion
5	1stand 2nd week of September	1st review Completed
6	3rd and 4th week of September	2nd review Completed
7	1st week of October	UML diagrams, State Charts and DFD's Completed
8	2nd week of October	3rd Review Completed
9	3rd and 4th week of October	Final review completed
10	November	Exam
11	December second week	Distribution of implementation modules
12	January	Project implementation in modules
13	February	Development of the entire module
14	March	Testing of the project
15	April first week	Final Report Submission

5.4 Team Organization

Team consists of four members. Proper planning mechanism is used and roles are analyzed and defined.

5.4.1 Team Structure

5.4.2 Management Reporting and Communication

Well organized plans were been made and completed accordingly within time. Progress reporting was been updated and completed. Communication as per requirements were being done effectively.

Table 5.4: Team Structure

SR No.	Member Name	Responsibility
1	Bezan Nanavatti	Developer , Project Lead and analysis, Documentation, Testing, Resource Gathering
2	Kavita Thakur	Developer , Requirement gathering, Documentation
3	Rutvik Korade	Developer , Project Design
4	Sunit Lohade	Developer , Testing

CHAPTER 6

PROJECT IMPLEMENTATION

6.1 Overview of Project Modules

1. Preprocessing:
   In this stage, the information which irrelevant for summarization or which can affect the final result is removed.

   (a) Data cleaning:

   - Tokenization:
     - Sentence tokenization: It means splitting the given text into sentences and storing them in a list. E.g.: "My name is Edward. I am a data scientist." After sentence tokenization: ['My name is Edward.', 'I am a data scientist.']
     - Word tokenization: It means splitting the given text into words and storing them in a list. E.g.: "My name is Edward." *After word tokenization:* ['My', 'name', 'is', 'Edward', '.']
   - Removing punctuations and extra white spaces: Removing punctuations is necessary because it holds no meaning value in the data and can also create and issue in differentiating words. Sometimes white spaces may also create problems in using the data so removing them is a good practice. E.g.: "Hi! My name is Edward." *After removing punctuations and extra white spaces:* Hi My name is Edward
   - Removing shortwords: The shortwords are words which contain apostrophe in them. So to convert them in into their normal form we perform this step. E.g.: "I've been working as data scientist for six years." *After removing shortwords:* I have been working as data scientist for six years.
   - Removing stopwords: The stopwords are words which are most common in any language. The words like he, she, and but, etc. are called stopwords. The use of these words does not hold any meaning to the data so they must be removed to reduce the feature. E.g.: "I am a data scientist." *After removing stopwords:* ['I', 'data', 'scientist', '.']

   (b) Normalization:

   - Lowercasing all the words: The lowercasing of data is a necessary step because languages like Python are case sensitive. So 'HELLO'

and 'hello' will be considered different. E.g.: "I am a data scientist." *After lowercasing all the words:* i am a data scientist.

- Lemmatization: It converts the word to it root form called lemma. This lemmatizing process, brings the word into its dictionary format. This is more accurate and makes analysis better. E.g.: studying, studies, etc. *After lemmatizing the word:* study

2. Sentiment Analysis:
   This is next step after pre-processing the data. In this stage, calculation of unique words, word count of sentences, geo-location, frequency of words will help us calculate the sentiment values or collective emotions of an area.

   - Twitter: On social media sites like Twitter, users actively engage in debate and share their opinions. Social media can be used to find out what people think and feel about certain events. There are numerous platforms for opinion-based data collecting and analytics that seek to elicit people's opinions on a range of topics.. There are multiple opinion-based data collection and analytics platforms that try to extract people's thoughts on various issues.
   - Implementation of Sentimental Analysis of Tweets: Reports on the tweets retrieved using the Twitter API, which is given by Twitter. Because of this Twitter API, there are a variety of tools for emotive or sentimental analysis of data on social media. This work made use of a number of readily available libraries.

3. Graph:
   A depressed interaction graph G is generated via some social graph model, minimizing the distance between the real and depressed interaction graphs. The input (actual) social media data is used to generate an interaction graph G. An interaction graph represents how social network users or actors interact with each other. In social media, entities and their interactions are recognized, and an interaction graph is constructed using a vertex set V for entities, an edge set E for interactions, and an attribute set A for both vertex (entity) and edge (interaction) attributes.

4. Post processing:
   This is the next step after processing. In this step, we use our collected sentiment values per area, divide them as such s = (negative sentiment

percentage)/(positive sentiment percentage) and then multiply this value with the number of crimes reported in that same area. We can call the result "count", which is used to create a simple output of a map which would display areas highly likely to witness crime.

CHAPTER 7

SOFTWARE TESTING

Software Testing is the process of executing every functionality and procedure of the program or application with the intent to find the errors or bugs. Testing is performed to investigate the entire project from every aspect. It deals with the motto to make the model more robust and accurate. The results make the developer aware of the issues that the program might go through in the future. Software testing is important to understand the future risks.

7.1 Type of Testing

7.1.1 Positive Testing

In this process the testing is done against valid data inputs. The model is being given the same kind of input which it is trained on and designed for. It only checks on the valid set of inputs. This checks whether the system actually shows error when the system is supposed to. Positive testing tries to predict whether the system shows the exact output as per the requirements and specifications. Its checks the expected behavior of the model.

7.1.2 Negative Testing

In this process the testing is done against invalid data inputs. The model is being given the different kinds of input which it is not trained on and not designed for accordingly. It only checks on the invalid set of inputs. This checks whether the system actually shows error when the system is supposed to. Negative testing tries to predict whether the system shows the exact output as per the requirements and specifications when the inputs are not in correspondence with the valid inputs. It checks the expected behavior of the model in case of random unwanted actions.

7.1.3 Unit Testing

The unit testing is performed on every part of the model that is being developed. It tests the model right from the stage wherein the input was been provided upto the state where the probabilistic labels are being generated. The tests test the inputs, outputs, functions, classes, modules and the entire data in chunks. This leads to the assurance of having no errors in the smallest possible module too. This helps in making the model more confident.

7.1.4 Integration Testing

In integration testing the entire system is being tested as a whole entity. The combination of various modules lead to the formation of an integrated system. The unit testing tests the individual elements and the integration testing tests all the unit combined as a single unit. It ensures that the functionality of the system is same and error free when combined completely.

7.2 Test cases & Test Results

Table 7.1: Test Cases and Test Results

Test Case ID	**Test Case Name**	**Objectives**	**Expected Output**	**Actual Output**	**Result**
1	Check Data Retrieval using Text functionality	To check whether the data is being collected	.csv file should be generated	File is generated	Pass
2	Check Data Integrity	To check whether the data is usable and in correct format	Data should be scanned	Data is Clean	Pass
3	Check if backend or code is generated and its working	To check functionality of code	Code should be working	Code Runs	Pass
4	Check final stage	To check if code has a functional UI	Code should have working user interface	program works	pass

CHAPTER 8

RESULTS

8.1 Screen-shots

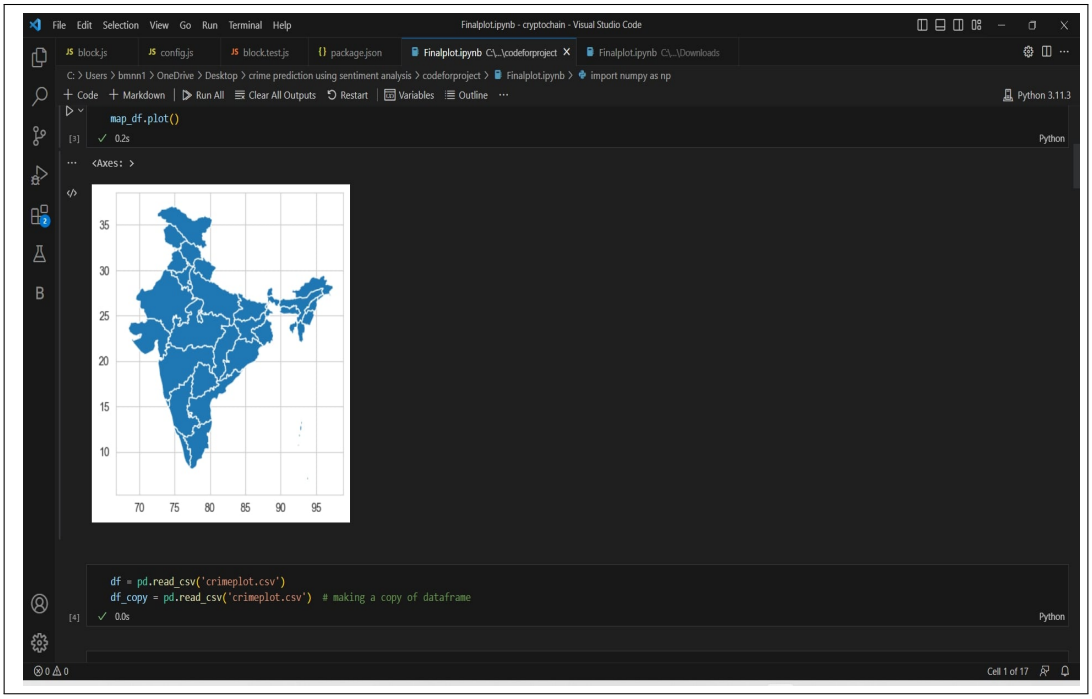


Figure 8.1: Map Structure [without plotting values]

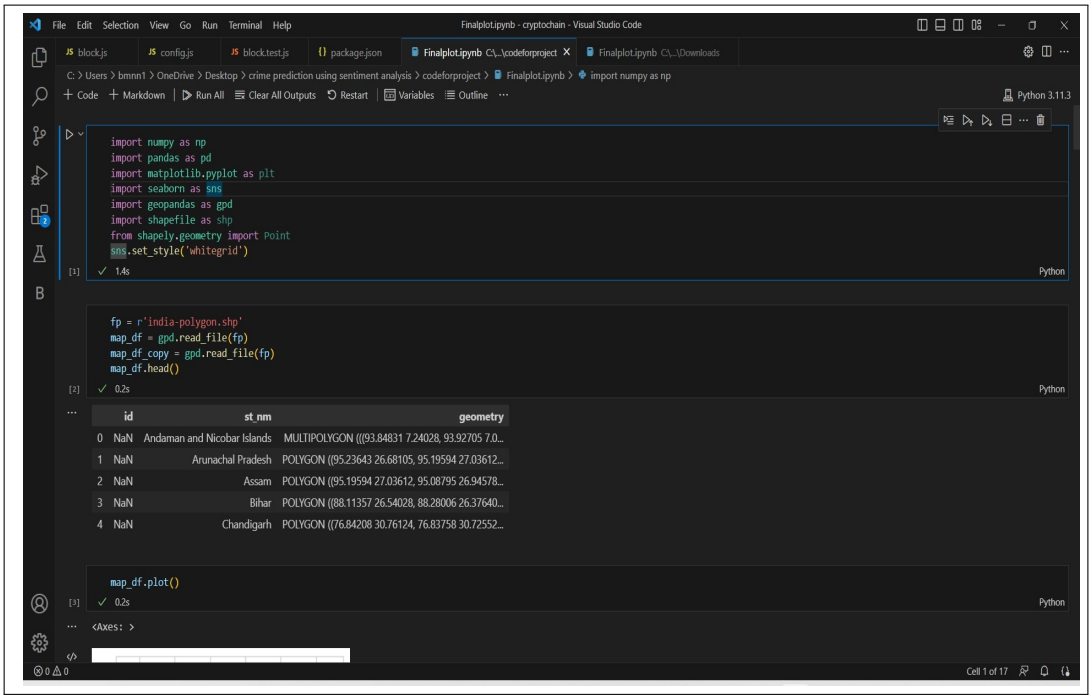


Figure 8.2: Main .exe libraries, map inherent data

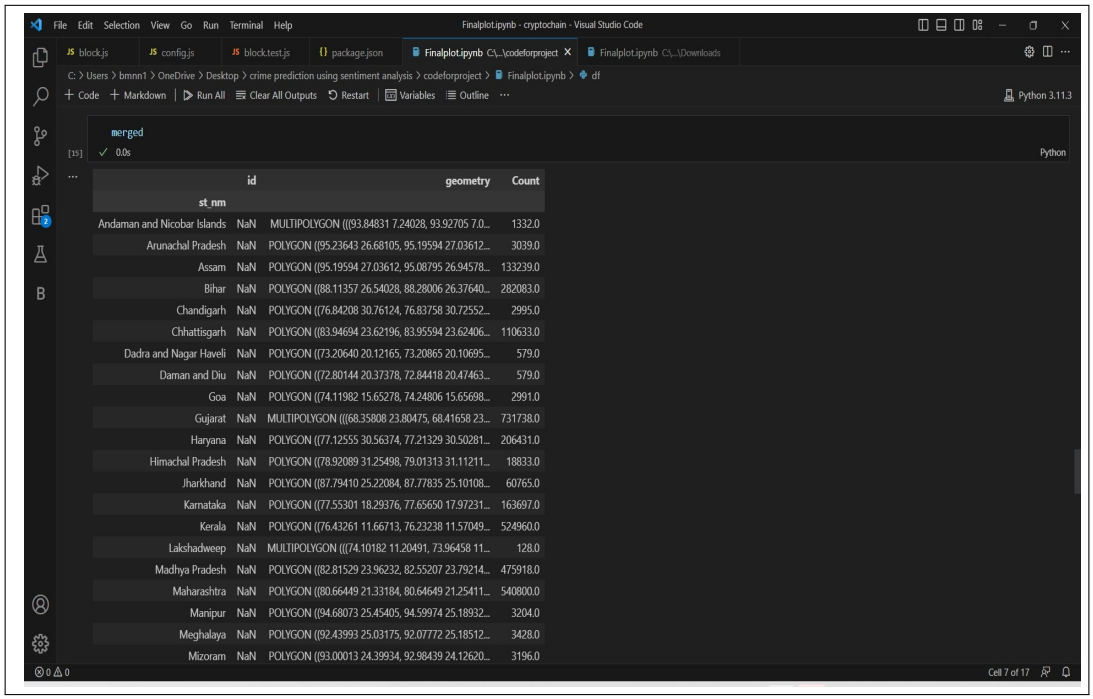


Figure 8.3: Merged data

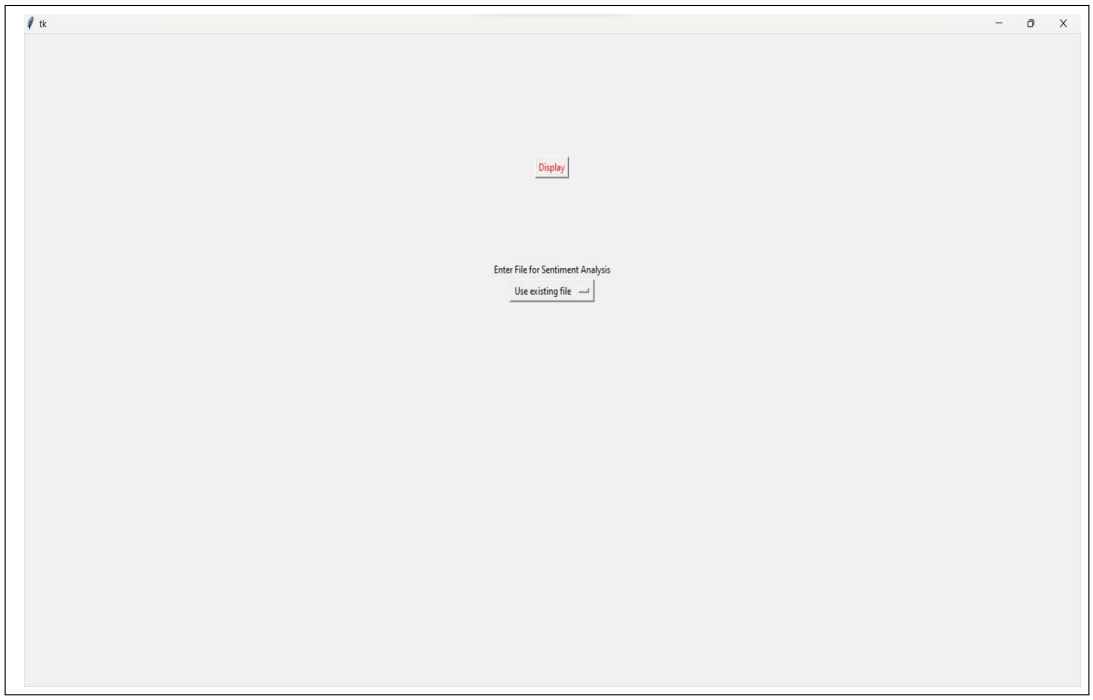


Figure 8.4: Opening page of our application

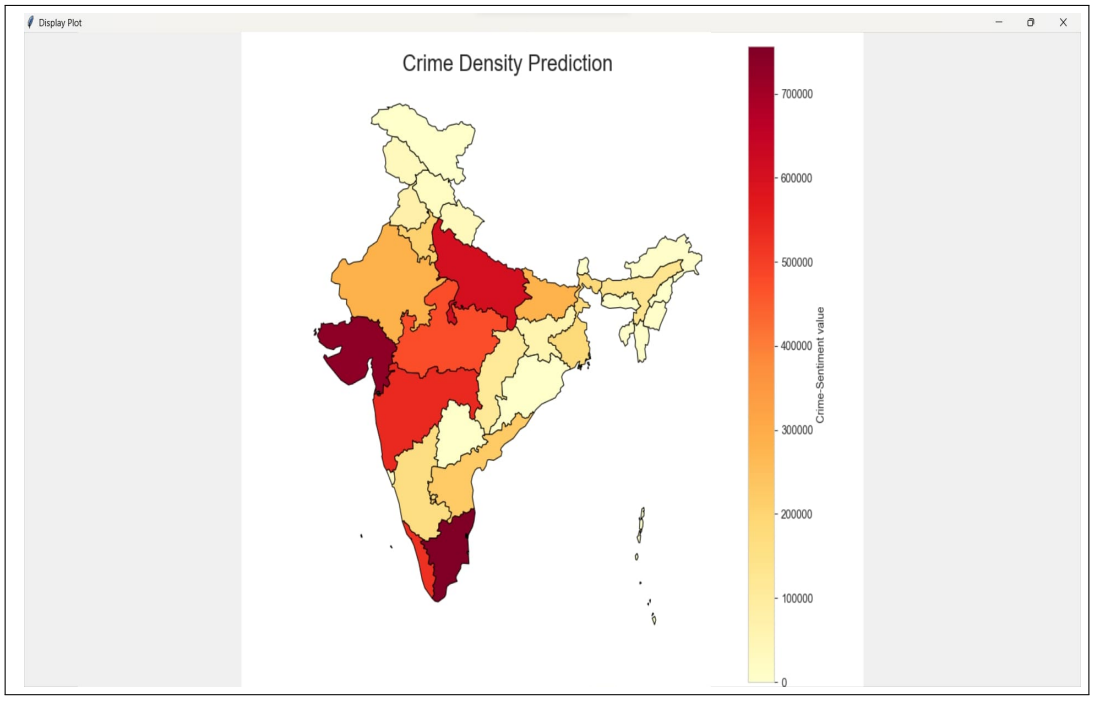


Figure 8.5: Final plot for our project code

CHAPTER 9

CONCLUSIONS AND FUTURE WORK

9.1 Conclusion

The aim is to predict the time [time variable will take time and more resources to implement which are not readily available for opensource use] and location in which a specific type of crime will occur. Our approach is based on sentiment analysis. Twitter is used extensively in the United States, India, The UK, as well as globally, creating many opportunities to augment decision support systems with Twitter-driven predictive analytics.

Twitter is an ideal data source for decision support: its users, who number in the millions, publicly discuss events, emotions, and innumerable other topics; its content is authored and distributed in real time at no charge; and individual messages (also known as tweets) are often tagged with precise spatial and temporal coordinates. The entire project is focused on prediction of crime using latest recent twitter data .More generally, this project has implications for decision makers concerned with geographic spaces occupied by Twitter-using individuals.

9.2 Future work / Applications

The research can be expanded upon to create profiles of significant influencers in a particular area of crime using psychological frameworks. The goal is to foresee when and where a particular kind of crime will occur. Our strategy is grounded in sentiment analysis. The main topics of this research's future work include a more in-depth semantic analysis of message content, temporal modelling, and the inclusion of auxiliary data sources. We may provide the data to state-level analysts to help them lower crime rates in their individual cities thanks to the effort that helps us identify crime rates. As a result of this work's effective crime detection capabilities, crime could be decreased in a number of cities. Future applications led by the government could use our softwares, as open-source modules to be able to efficiently interject and possibly interrupt the criminal intents nearing an area The research can be further developed to form profiles of important influencers in a specific topic related to crime using psychological frameworks. The aim is to predict the time and location in which a specific type of crime will occur. Our approach is based on sentiment analysis The important areas of future work for this research, including deeper semantic analysis of message content, temporal modeling, and incorporation of auxiliary data sources.

9.3 BIBLIOGRAPHY:

Pratibha, A. Gahalot, Uprant, S. Dhiman and L. Chouhan, "Crime Prediction and Analysis," 2nd International Conference on Data, Engineering and Applications (IDEA), 2020, pp. 1-6, doi: 10.1109/IDEA49133.2020.9170731.

[2] Kumar, A. Verma, G. Shinde, Y. Sukhdeve and N. Lal, "Crime Prediction Using K-Nearest Neighboring Algorithm," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-4, doi: 10.1109/ic-ETITE47903.2020.155.

[3] N. H. M. Shamsuddin, N. A. Ali and R. Alwee, "An overview on crime prediction methods," 2017 6th ICT International Student Project Conference (ICT-ISPC), 2017, pp. 1-5, doi: 10.1109/ICT-ISPC.2017.8075335.

[4] W. Safat, S. Asghar and S. A. Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques," in IEEE Access, vol. 9, pp. 70080-70094, 2021, doi:10.1109/ACCESS.2021.3078117.

[5] S. G. Krishnendu, P. P. Lakshmi and L. Nitha, "Crime Analysis and Prediction using Optimized K-Means Algorithm," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 915-918, doi: 10.1109/ICCMC48092.2020.ICCMC-000169.

[6] J. Kiran and K. Kaishveen., "Prediction Analysis of Crime in India Using a Hybrid Clustering Approach," 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference on, 2018, pp. 520-523, doi: 10.1109/I-SMAC.2018.8653744.

[7] X. Chen, Y. Cho and S. Y. Jang, "Crime prediction using Twitter sentiment and weather," 2015 Systems and Information Engineering Design Symposium, 2015, pp. 63-68, doi: 10.1109/SIEDS.2015.7117012.

M. Al Boni and M. S. Gerber, "Predicting crime with routine activity patterns inferred from social media," 2016 IEEE International Conference on Systems,

APPENDIX A

1. **Statistical and Analytical proceeding:**
   This appendix reveals that our project has used the following principles and methods to achieve the final result/goal.

   Methods:

   Random forest, Decision trees, Support Vector Mechanism/Machine, Natural Language Processing, User Interface Creation.

2. **The project has aimed and acomplished:**

   [i] Systematic and Organised Data Collection.

   [ii] Functional Backend/Code using Python.

   [iii] Functional User Interface to display our results.

APPENDIX B

- **Conference Proceeding :**
  INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT), 2023

- **Conference Date :**

  21, May 2023

- **Held by :**

  INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT), PUNE, MAHARASHTRA

IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CRIME DETECTION USING SENTIMENT ANALYSIS

Bezan Nanavatti, Kavita Thakur, Rutvik Korade, Sunit Lohade
ALL INDIA SHRI SHIVAJI MEMORIAL SOCIETY
INSTITUTE OF INFORMATION TECHNOLOGY
PUNE, MAHARASHTRA

Abstract: Crime is one of the biggest, and most pervasive problems in our society, and preventing it is a critical duty. Every day, many crimes are committed in large numbers. This makes it an absolute necessity to monitor and make note of all crimes. A database for them that can be used as a resource in the future. In order to predict and solve future crimes, it is currently difficult to keep proper records of criminal activity. The goal of this research is to examine a dataset that includes values of a large number of crimes, in correspondence with sentiment values, and make predictions about intensity and chance of crimes that may occur.

I. INTRODUCTION

The greatest problem, or in a sense, tragedy, to befall humanity is crime. Crimes occur on a regular basis and is a menacing factor in almost every human settlement and environment. To add to this, the amount only increases exponentially. Little towns, large cities, and rural areas all experience crime. Due to this ungodly rise in crime, cases must be resolved at a significantly quicker pace.

We can, at times, anticipate the type and/or intensity of crime[s] that might occur in a specific location, by utilizing a machine learning system with past data or sentiment or some arithmetic logic as it's preface.

II. LITERATURE SURVEY

- Pratibha, A. Gahalot, Uprant, S. Dhiman and L. Chouhan, "Crime Prediction and Analysis," 2nd International Conference on Data, Engineering and Applications (IDEA),
- Kumar, A. Verma, G. Shinde, Y. Sukhdeve and N. Lal, "Crime Prediction Using K-Nearest Neighboring Algorithm," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE),
- N. H. M. Shamsuddin, N. A. Ali and R. Alwee, "An overview on crime prediction methods," 2017 6th ICT International Student Project Conference (ICT-ISPC)
- W. Safat, S. Asghar and S. A. Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques,"
- J. Kiran and K. Kaishveen., "Prediction Analysis of Crime in India Using a Hybrid Clustering Approach,"
- M. Al Boni and M. S. Gerber, "Predicting crime with routine activity patterns inferred from social media,"

- Gerber, Matthew. (2014). Predicting Crime Using
- Twitter and Kernel Density Estimation, Decision Support Systems.
- X. Chen, Y. Cho and S. Y. Jang, "Crime prediction using Twitter sentiment and weather," 2015 Systems and Information Engineering Design Symposium

III. PROPOSED WORK

The ideal goal and aim of this work is to classify places or areas with the largest number of crimes by applying sentiment analysis provided by an open source dataset of information, namely, “Twitter messages”, to estimate crime rate or density of the same. By giving data on crime at various and all levels in terms of area and location, this work will prove to be useful in aiding an envisioned fall or drop, in crime rates.

Multiple unnerving incidents have left people as a whole, feeling devastatingly fearful for the rest of their lives, maybe because of that one time when they were made to do something they didn't want to do [an example]. Our work's major goal is to primarily concentrate or focus on the functionality and accessibility of social media in promoting and hopefully ensuring the safety of everyone in the world, with a particular reference to the accessibility of social-media website and application-- Twitter.

IV. METHODOLOGY

TWITTER ANALYSIS

Social media sites like Twitter encourage users to actively participate in discussions and share their opinions. Social media can be used to find out what people think and feel about a particular event. There are many platforms for collecting and analyzing opinion-based data that aim to get people's opinions on various topics. There are numerous opinion-based data collection and analysis platforms that attempt to elicit people's thoughts on various topics. Implementation of Sentimental Analysis of Tweets helps create Reports on the tweets retrieved using the Twitter API, which is given by Twitter. Because of this Twitter API, there are a variety of tools for emotive or sentimental-analysis of data on social media. This work made use of a number of readily available libraries, and all data sources and values are open sourced and not copy-righted.

GRAPH

An interactive graph is generated via some social graph model, to introduce a visual representation of the output data of our project. The input would refer to social media data, like data from tweets. An interaction graph represents how social network users interact with each other. In social media, entities [or users] and their interactions are recognized, and an interaction graph is constructed using a vertex set for entities, an edge set for interactions, and an attribute set for both, the vertex (entity) and edge (interaction) attributes.

FINAL REPORT SUMMARIZATION

If there are disproportionately many or purely neutral tweets, it means that people are actually less engaged in the topic at hand, and are unwilling to take sides. Although people's attitudes can change depending on the circumstances, it is also important to keep in mind that the experiments or project’s results are susceptible to change depending on the data. For instance, rape news becomes the most popular topic of the year in 2020. The fact that neutral tweets make up more than 66% of some inquiries makes the viewpoints plainly limited. Considering an example of an analysis performed in relevance to "How safe some areas are", we could conclude that one area is far safer than another.

V. ALGORITHM

Support Vector Machine (SVM) is an administered, AI calculation is included and can be utilized for both grouping and relapse difficulties. Nonetheless, it is generally utilized in order to prevent computational errors and issues. With respect to this calculation, each piece of information is represented as a point in an multi-dimensional space with the value of each component being the value of a certain arrangement. When that happens, we can execute grouping by locating the hyper-plane that clearly divides the groups.

The coordinates of a single perception are essentially what support vectors are. Support Vector Machine is the method that isolates the two classes (hyper-plane/line) the best. A help vector machine constructs a hyper plane or collection of hyper planes in a high- or infinite dimensional space, which can be used for grouping, relapsing, or other tasks like anomalies location. Intuitively, a good partition is produced by the hyper plane with the greatest distance from the nearest prepared information point of any class (referred to as the practical edge), as, generally speaking, the larger the edge, the fewer errors the classifier makes in its predictions. It commonly happens that the sets to separate are not clearly distinguishable in that area, despite the fact that the first problem may be described in a small dimensional space.

VI. ARCHITECTURE

The user signs in via a home page or opens a local application to access our software, after which he/she uploads the data (or tweets), and then implements sentimental analysis of comments or values stored in the database in the form of positive and negative text analysis, based on which a modified graph is formed. To add to this, we can use available data on crimes committed along with the sentiment of the area to provide our uses with a map entailing the crime as well as sentiment in a proportionally defined map with colors or indicators which describe the “best-of” or “worst-of” both factors (crimes reported and sentiment value of the same area).

An example of sentiment analysis in a tabular format is given following this paragraph, with the positive and the negative sentiment values loaded in percentages over the total tweets in the area. (Tentative values are used here)

Table- Analysis in tabular format:

S. No.	Name of Cities	Positive Analysis (%)	Negative Analysis (%)
1	Lucknow	9	1
2	Delhi	14	20
3	Kolkata	9	11
4	Pune	12	8
5	Mumbai	18	2
6	Goa	9	11
7	Chennai	13	9

A diagrammatic format which represents the working of our application is presented as follows:

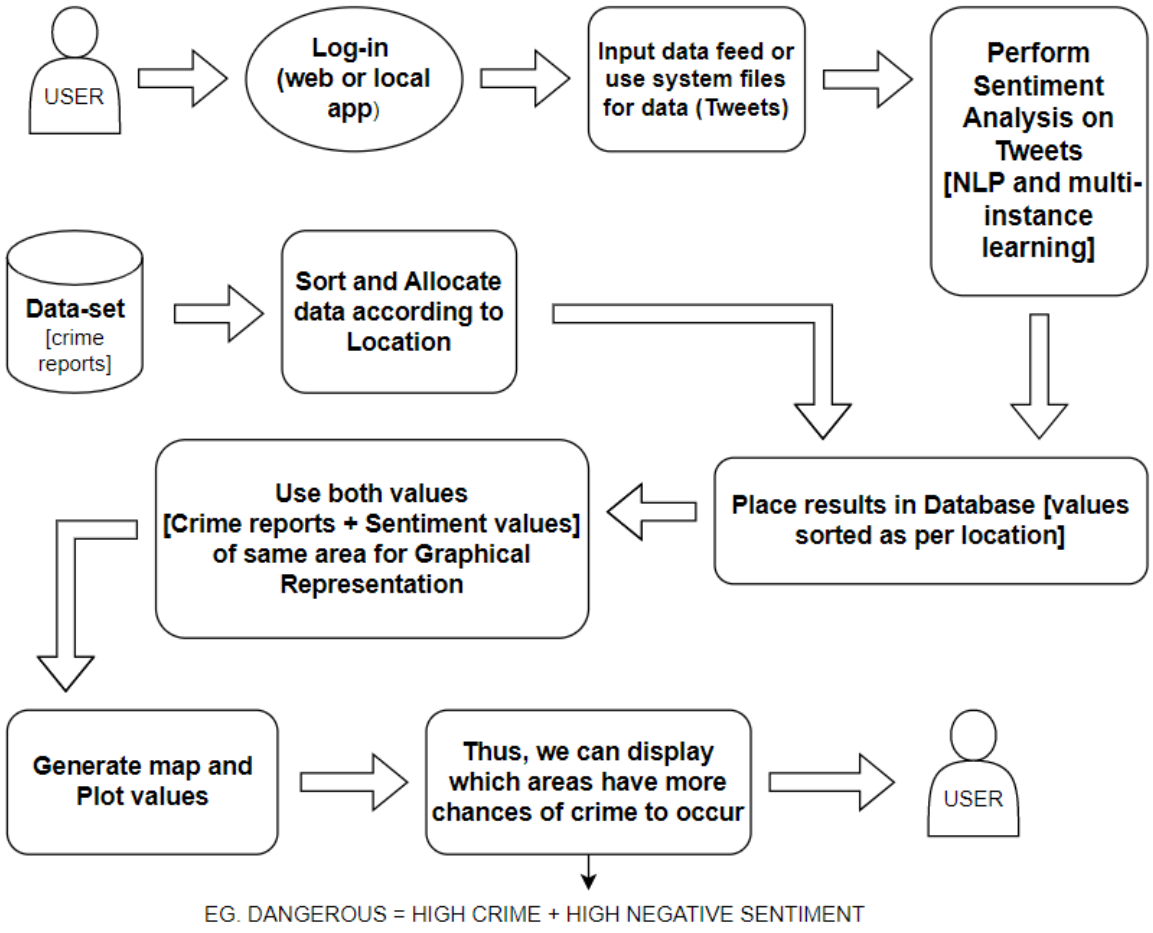


Figure 1. Architecture/Working of our system

VII. CONCLUSION

In this study, we examined a number of AI evaluations that can assist us in sorting through and breaking down the enormous amount of Twitter data we have gathered, which includes a significant number of tweets and instant messages released every day. When it comes to analyzing massive volumes of data, certain AI calculations like the SPC computation and direct logarithmic Factor Model techniques are incredibly compelling and helpful. They also aid in grouping the data into pertinent categories. A different type of AI calculation called a backing vector machine is frequently used to extract crucial information from Twitter and create a structured inlay of data values that can be used in our analysis and arithmetic algorithms to help approach our target, or aim, which is to predict the time and/or location in which a specific type of crime will occur. Our approach is based on sentiment analysis. Twitter is used extensively in the United States as well as globally, creating many opportunities to augment decision support systems with Twitter-driven predictive analytics. Twitter is an ideal data source for decision support: its users, who number in the millions, publicly discuss events, emotions, and innumerable other topics; its content is authored and distributed in real time at no charge; and individual messages (also known as tweets) are often tagged with precise spatial and temporal coordinates.

VIII. FUTURE SCOPE

This research can be expanded upon to create profiles of significant influence in particular areas of crime using multiple psychological frameworks. The future goal is to foresee when and where a particular kind of crime will occur. Our strategy is grounded in (dedicated to) sentiment analysis. The main topics of this research's future work include a more in-depth analysis of message content, temporal modelling, and the inclusion of auxiliary data sources. We may provide the data to state-level analysts to help them lower crime rates in their individual cities thanks to the effort that helps us identify crime rates, and create an open-source platform for all to access it. As a result of this work's effective crime detection capabilities, crime could be decreased in a number of cities. Future programs or procedures led by the government can implement our analysis methods and software, etc. to aid them in preparation of crimes that may occur and information analysis and collection of data that may make it easier for them to tackle it.

IX. REFERENCES

1) >Agarwal, S., Godbole, S., Punjani, D., & Roy, S.(2007, October). How much noise is too much: A study in automatic text classification. In Seventh IEEE International Conference on Data Mining (ICDM 2007) (pp. 3-12). IEEE.
2) >Ahmad, K., Han, J., Hutson, E., Kearney, C., & Liu, S. (2016). Media-expressed negative tone and firmlevel stock returns. Journal of Corporate Finance, 37, 152-172.
3) >Baid, P., Gupta, A., & Chaplot, N. (2017). Sentiment analysis of movie reviews using machine learning techniques. International Journal of Computer Applications, 179(7), 45-49. >Bollen, J., Mao, H., & Pepe, A. (2011, July).Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In Proceedings of the international AAAI conference on web and social media (Vol. 5, No. 1).
4) >Choudhury, M., Saraf, R., Jain, V., Mukherjee, A., Sarkar, S., & Basu, A. (2007). Investigation and modeling of the structure of texting language. International Journal of Document Analysis and Recognition (IJDAR), 10(3), 157-174.
5) >Diakopoulos, N. A., & Shamma, D. A. (2010, April). Characterizing debate performance via aggregated twitter sentiment. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 1195-1198).
6) >Gamon, M. (2004). Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. In COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics (pp. 841-847).
7) >Kabir, M., Kabir, M. M. J., Xu, S., & Badhon, B. (2019). An empirical research on sentiment analysis using machine learning approaches. International Journal of Computers and Applications, 1-9.
8) >Kolchyna, O., Souza, T. T., Treleaven, P., & Aste, T. (2015). Twitter sentiment analysis: Lexicon method, machine learning method and their combination. arXiv preprint arXiv:1507.00955.
9) >Matsumoto, S., Takamura, H., & Okumura, M. (2005,May). Sentiment classification using word subsequences and dependency sub-trees. In Pacific-Asia conference on knowledge discovery and data mining (pp. 301-311). Springer, Berlin, Heidelberg. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey.
10) >Ain Shams engineering journal, 5(4), 1093-1113. O'Hare, N., Davy, M., Bermingham, A., Ferguson, P., Sheridan, P., Gurrin, C., & Smeaton, A. F. (2009, November). Topic-dependent sentiment analysis of financial blogs. In Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion (pp. 9-16).
11) >Pang, B., & Lee, L. (2009). Opinion mining and sentiment analysis. Comput. Linguist, 35(2), 311-312. Renault, T. (2020). Sentiment analysis and machine learning in finance: a comparison of methods and models on one million messages. Digital Finance, 2(1), 1-13.
12) >Sarlan, A., Nadam, C., & Basri, S. (2014, November).Twitter sentiment analysis. In Proceedings of the 6th International conference on Information Technology and Multimedia (pp. 212-216). IEEE.
13) >Siddiqui, S., Darbari, M., & Yagyasen, D. (2020). Modelling and Simulation of Queuing Models through the concept of Petri Nets. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 9(3), 17-28.
14) >Wang, S. I., & Manning, C. D. (2012, July). Baselines and bigrams: Simple, good sentiment and topic
15) >And multiple open-sourced data and information aided in creating a path to success for our undertaken project here.

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

Kavita Thakur

In recognition of the publication of the paper entitled

CRIME DETECTION USING SENTIMENT ANALYSIS

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 5 May 2023 , Date of Publication: 19-May-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2305846

Registration ID : 237431

ISSN 2320-2882 IJCRT

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

Bezan Nanavatti

In recognition of the publication of the paper entitled

CRIME DETECTION USING SENTIMENT ANALYSIS

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 5 May 2023 , Date of Publication: 19-May-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2305846

Registration ID : 237431

ISSN 2320-2882 IJCRT

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

Sunit Lohade

In recognition of the publication of the paper entitled

CRIME DETECTION USING SENTIMENT ANALYSIS

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 5 May 2023 , Date of Publication: 19-May-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2305846

Registration ID : 237431

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

Rutvik Korade

In recognition of the publication of the paper entitled

CRIME DETECTION USING SENTIMENT ANALYSIS

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 5 May 2023 , Date of Publication: 19-May-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2305846

Registration ID : 237431

ISSN 2320-2882 IJCRT

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

Certificate of Publication

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | ISSN: 2320 - 2882

An International Open Access, Peer-reviewed, Refereed Journal

The Board of

International Journal of Creative Research Thoughts

Is hereby awarding this certificate to

Minal Zope

In recognition of the publication of the paper entitled

CRIME DETECTION USING SENTIMENT ANALYSIS

Published In IJCRT (www.ijcrt.org) & 7.97 Impact Factor by Google Scholar

Volume 11 Issue 5 May 2023 , Date of Publication: 19-May-2023

UGC Approved Journal No: 49023 (18)

PAPER ID : IJCRT2305846

Registration ID : 237431

ISSN 2320-2882 IJCRT

EDITOR IN CHIEF

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly Journal

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS | IJCRT

An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal

Website: www.ijcrt.org | Email id: editor@ijcrt.org | ESTD: 2013

IJCRT | ISSN: 2320-2882 | IJCRT.ORG

IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Ref No : IJCRT/Vol 11/ Issue 5 / 846

To,
Kavita Thakur

Subject: Publication of paper at International Journal of Creative Research Thoughts.

Dear Author,

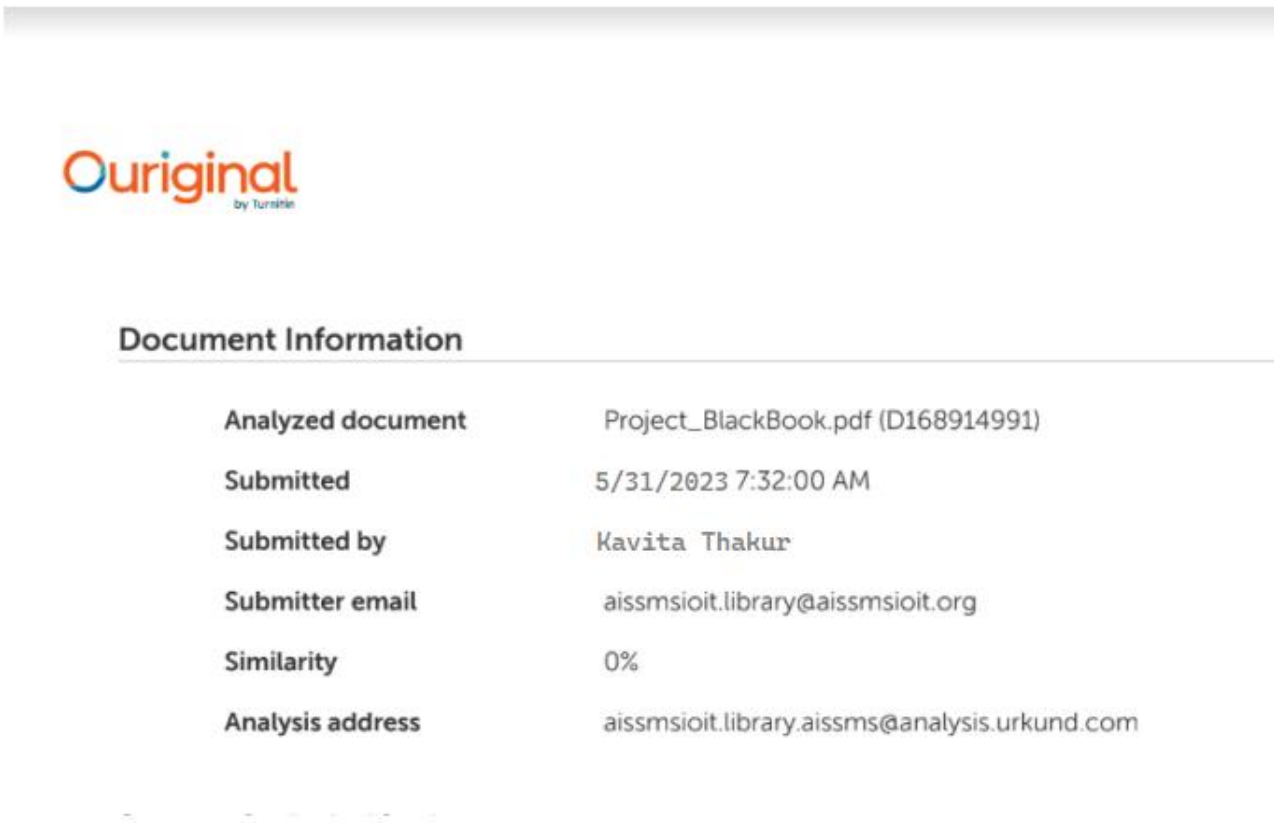
With Greetings we are informing you that your paper has been successfully published in the International Journal of Creative Research Thoughts - IJCRT (ISSN: 2320-2882). Thank you very much for your patience and cooperation during the submission of paper to final publication Process. It gives me immense pleasure to send the certificate of publication in our Journal. Following are the details regarding the published paper.

About IJCRT	:	Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly, Indexing in all major database & Metadata, Citation Generator, Digital Object Identifier(DOI) | UGC Approved Journal No: 49023 (18)
Registration ID	:	IJCRT_237431
Paper ID	:	IJCRT2305846
Title of Paper	:	CRIME DETECTION USING SENTIMENT ANALYSIS
Impact Factor	:	7.97 (Calculate by Google Scholar) | License by Creative Common 3.0
Publication Date	:	19-May-2023
DOI	:	
Published in	:	Volume 11 | Issue 5 | May 2023
Page No	:	g930-g934
Published URL	:	http://www.ijcrt.org/viewfull.php?&p_id=IJCRT2305846
Authors	:	Kavita Thakur, Bezan Nanavatti, Sunit Lohade, Rutvik Korade, Minal Zope
Notification	:	UGC Approved Journal No: 49023 (18)

Thank you very much for publishing your article in IJCRT.

Editor In Chief
International Journal of Creative Research Thoughts - IJCRT
(ISSN: 2320-2882)

APPENDIX C: PLAGIARISM



Ouriginal
by Turnitin

Document Information

Analyzed document	Project_BlackBook.pdf (D168914991)
Submitted	5/31/2023 7:32:00 AM
Submitted by	Kavita Thakur
Submitter email	aissmsioit.library@aissmsioit.org
Similarity	0%
Analysis address	aissmsioit.library.aissms@analysis.urkund.com

AISSMS
INSTITUTE OF INFORMATION TECHNOLOGY
ADDING VALUE TO ENGINEERING

Department of Computer Engineering

Self-Evaluation

Name of the Project	Crime Prediction Using Sentiment Analysis
Name of the student	Bezan Nanavatti
Name of the student	Kavita Thakur
Name of the student	Rutvik Korade
Name of the student	Sunit Lohade

File of literature Survey	Design	Implementation	Task and Result	Attendance on project Day	Work according to plan Activity	Managing of log book	Project Presentation or Participation	Project exhibition participation	Award Prize if any
(5)	(20)	(20)	(20)	(5)	(10)	(5)	(5)	(5)	(5)

Observations and comments of Guide:

Name of Student **Signature of student** **Signature of Guide**

1. Bezan Nanavatti

2. Kavita Thakur

3. Rutvik Korade

4. Sunit Lohade